Lösningsförslag till
Matematisk statistik och sannolikhetslära
2020-08-26

1. (a) Per definition är

$$P(B \mid A \cup B) = \frac{P(B \cap (A \cup B))}{P(A \cup B)} = \frac{P(B)}{P(A \cup B)},$$

så

$$P(B) = P(A \cup B) \cdot P(B \mid A \cup B) = 0.8 \cdot 0.75 = 0.6,$$

så

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0.65 + 0.6 - 0.8 = 0.45$$

medan

$$P(A)P(B) = 0.65 \cdot 0.6 = 0.39 \neq 0.45,$$

så A och B är inte oberoende.

(b) Vi får

$$\frac{13 \cdot \binom{4}{3} \cdot 12 \cdot \binom{4}{2}}{\binom{52}{5}} = \frac{3744}{2\,598\,960} \approx \frac{1}{694} \quad \text{eller} \quad 0.144\ \%,$$

där 13 och $\binom{4}{3}$ är valet av valör respektive färger på trissen, och 12 och $\binom{4}{2}$ är valet av valör respektive färger på paret.

(c) Vi får

$$\frac{4 \cdot \left(\binom{13}{5} - 10\right)}{\binom{52}{5}} = \frac{5108}{2\,598\,960} \approx \frac{1}{509} \quad \text{eller} \quad 0.197\ \%,$$

där 4 är valet av färg, $\binom{13}{5}$ är valet av valörer och 10 är valet av minsta valör i följden (som inte kan vara knekt, dam eller kung).

Svar: (a) $P(A \cap B) = 0.45$; A och B är inte oberoende, (b) 1/694, (c) 1/509.

2. (a) Vi har

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy = \int_x^{\infty} e^{-y}\,dy = \Big[-e^{-y}\Big]_{y=x}^{\infty} = e^{-x}, \quad x \geq 0,$$

och

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx = \int_0^{y} e^{-y}\,dx = \Big[xe^{-y}\Big]_{x=0}^{y} = ye^{-y}, \quad y \geq 0.$$

(Notera att $X \in \mathrm{Exp}(1)$ och $Y \in \Gamma(2,1)$.)

(b) Rita figur! Vi har

$$P(X > 2,\, Y < 4) = \int_2^4 \int_2^y e^{-y}\,dx\,dy = e^{-2} - 3e^{-4}$$

eller

$$P(X > 2,\, Y < 4) = \int_2^4 \int_x^4 e^{-y}\, dy\, dx = e^{-2} - 3e^{-4}.$$

Vidare är

$$P(Y < 4) = \int_{-\infty}^4 f_Y(y)\, dy = \int_0^4 ye^{-y}\, dy = 1 - 5e^{-4},$$

så

$$P(X > 2 \,|\, Y < 4) = \frac{P(X > 2,\, Y < 4)}{P(Y < 4)} = \frac{e^{-2} - 3e^{-4}}{1 - 5e^{-4}} = \frac{e^2 - 3}{e^4 - 5} \approx 0.0885.$$

Svar: (a) $f_X(x) = e^{-x}$, $x \geq 0$; $f_Y(y) = ye^{-y}$, $y \geq 0$, (b) $P(X > 2 \,|\, Y < 4) = \dfrac{e^2 - 3}{e^4 - 5}$.

3. (a) Sats 3.4 och binomialsatsen ger

$$g(t) = Et^X = \sum_{k=0}^{n} t^k \binom{n}{k} p^k q^{n-k} = \sum_{k=0}^{n} \binom{n}{k} (pt)^k q^{n-k} = (pt + q)^n.$$

(b) Vi får

$$g'(t) = n(pt + q)^{n-1} \cdot p$$

och

$$g''(t) = n(n-1)(pt + q)^{n-2} \cdot p^2,$$

så $EX = g'(1) = np$ och $g''(1) = n(n-1)p^2$, ty $p + q = 1$, så

$$\begin{aligned} VX &= g''(1) + g'(1) - (g'(1))^2 = n(n-1)p^2 + np - (np)^2 = n^2p^2 - np^2 + np - n^2p^2 \\ &= np - np^2 = np(1-p). \end{aligned}$$

4. (a) Låt X beteckna antalet godkända tentander. Då är $X \in \mathrm{Bin}(60, 0.65)$, så $VX = 13.65 > 5$, så CGS medför att $X \approx \mathrm{N}(39, 13.65)$, så med halvkorrektion fås

$$\begin{aligned} P(X \geq 42) &= 1 - P(X \leq 41.5) = 1 - P\left(\frac{X - 39}{\sqrt{13.65}} \leq \frac{41.5 - 39}{\sqrt{13.65}}\right) \approx 1 - \Phi(2.5/\sqrt{13.65}) \\ &\approx 1 - \Phi(0.68) \approx 1 - 0.7517 = 0.2483. \end{aligned}$$

Halvkorrektion krävs för full poäng. Utan halvkorrektion fås

$$\begin{aligned} P(X \geq 42) &= 1 - P(X \leq 41) = 1 - P\left(\frac{X - 39}{\sqrt{13.65}} \leq \frac{41 - 39}{\sqrt{13.65}}\right) \approx 1 - \Phi(2/\sqrt{13.65}) \\ &\approx 1 - \Phi(0.54) \approx 1 - 0.7054 = 0.2946. \end{aligned}$$

Som jämförelse fås $P(X \geq 42) \approx 0.2518$ utan normalapproximation.

(b) Låt Y och X beteckna antalet studenter som anmäler sig till respektive fullföljer kursen. Då är $Y \in \mathrm{Po}(\lambda)$ och $(X \mid Y = n) \in \mathrm{Bin}(n, p)$, så lagen om total sannolikhet ger

$$
\begin{aligned}
P(X = j) &= \sum_{k=0}^{\infty} P(Y = k)P(X = j \mid Y = k) = \sum_{k=j}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} \cdot \binom{k}{j} p^j q^{k-j} \\
&= \frac{(\lambda p)^j}{j!} e^{-\lambda} \sum_{k=j}^{\infty} \frac{(\lambda q)^{k-j}}{(k-j)!} = \frac{(\lambda p)^j}{j!} e^{-\lambda} e^{\lambda q} = \frac{(\lambda p)^j}{j!} e^{-\lambda p},
\end{aligned}
$$

dvs. $X \in \mathrm{Po}(\lambda p)$.

Svar: (a) 24.8 %.

5. (a) Låt X och Y beteckna antalet lyckade försök för Ada respektive antalet försök för Beda. Då är $X \in \mathrm{Bin}(8, p)$ och $Y \in \mathrm{NegBin}(2, p)$, så likelihoodfunktionen

$$
\begin{aligned}
L(p) = p_1(x; p) \cdot p_2(y; p) &= \binom{8}{x} p^x (1-p)^{8-x} \cdot \binom{y-1}{2-1} p^2 (1-p)^{y-2} \\
&= \binom{8}{3} p^3 (1-p)^5 \cdot 6p^2 (1-p)^5 = 336 p^5 (1-p)^{10},
\end{aligned}
$$

så

$$\ln L(p) = \ln 336 + 5 \ln p + 10 \ln(1-p),$$

så

$$\frac{d}{dp}\Big(\ln L(p)\Big) = \frac{5}{p} - \frac{10}{1-p} = \frac{5 - 15p}{p(1-p)} = 0 \iff p = 1/3,$$

som ger maximum. Alltså är ML-skattningen $p^* = 1/3$.

(b) Låt x_i och y_i beteckna deltagare i:s före/efter-resultat. Vi antar att x_i är en observation av $X_i \in \mathrm{N}(\mu_i, \sigma_1^2)$ och att y_i är en observation av $Y_i \in \mathrm{N}(\mu_i + \Delta, \sigma_2^2)$, där alla X_i och Y_i är oberoende. Sätt $z_i = y_i - x_i$. Då är z_i en observation av $Z_i \in \mathrm{N}(\Delta, \sigma^2)$, där $\sigma^2 = \sigma_1^2 + \sigma_2^2$, så $\bar{Z} \in \mathrm{N}(\Delta, \sigma^2/5)$, så

$$\frac{\bar{Z} - \Delta}{s_z/\sqrt{5}} \in t(5-1) = t(4).$$

Vi ska testa H_0: $\Delta = 0$ mot H_1: $\Delta > 0$. Som testvariabel väljer vi

$$T = T(z) = \frac{\bar{z}}{s_z/\sqrt{5}},$$

som alltså är en observation av $T(Z) \in t(4)$ om H_0 gäller. Med H_1: $\Delta > 0$ blir det kritiska området $C = \{T \geq t_{0.05}(4)\}$. Vi får $\bar{z} = 9$ och $s_z^2 = 92.5$, så $T \approx 2.09 < 2.13 \approx t_{0.05}(4)$, så H_0 accepteras, dvs. tränaren har inte fått ett signifikant resultat vid 5 % felrisk.

Alternativ lösning: Det ensidiga konfidensintervallet är $(-0.17, \infty)$, som innehåller 0, så H_0 accepteras.

Svar: (a) $p^* = 1/3$, (b) nej, det har hon inte.

6. (a) Låt X beteckna vikten av en pumpa. Då är

$$\frac{(6-1)s^2(X)}{\sigma^2} \in \chi^2(6-1) = \chi^2(5),$$

vilket ger konfidensintervallet

$$I_{\sigma^2} = \left(\frac{5s^2}{\chi^2_{0.025}(5)}, \frac{5s^2}{\chi^2_{0.975}(5)} \right) = \left(\frac{S_{xx}}{\chi^2_{0.025}(5)}, \frac{S_{xx}}{\chi^2_{0.975}(5)} \right) \approx (0.89, 13.76),$$

ty $\chi^2_{0.025}(5) \approx 12.833$, $\chi^2_{0.975}(5) \approx 0.8312$ och

$$S_{xx} = \sum_{i=1}^{6} (x_i - \bar{x})^2 = \sum_{i=1}^{6} x_i^2 - 6\bar{x}^2 = 11.44,$$

så

$$I_\sigma \approx \left(\sqrt{0.89}, \sqrt{13.76} \right) \approx (0.94, 3.71).$$

(b) Vi har $X \in \mathrm{N}(\mu, \sigma^2)$, så $\bar{X} \in \mathrm{N}(\mu, \sigma^2/n)$, så

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \in \mathrm{N}(0, 1).$$

Vi ska testa H_0: $\mu = 9.6$ mot H_1: $\mu < 9.6$. Som testvariabel väljer vi

$$T = T(x) = \frac{\bar{x} - 9.6}{1.5/\sqrt{6}},$$

som alltså är en observation av $T(X) \in \mathrm{N}(0, 1)$ om H_0 gäller. Med H_1: $\mu < 9.6$ blir det kritiska området $C = \{T \leq -\lambda_{0.01}\}$. Vi får $\bar{x} = 8.1$, så $T \approx -2.45 \leq -2.33 \approx -\lambda_{0.01}$, så H_0 förkastas till förmån för H_1, dvs. $\mu < 9.6$ vid 1 % felrisk.

Alternativ lösning 1: Det ensidiga konfidensintervallet är $(0, 9.52)$, som inte innehåller 9.6, så H_0 förkastas till förmån för H_1.

Alternativ lösning 2: P-värdet är

$$\begin{aligned} P &\overset{\text{defn}}{=} P_{H_0}(\text{att få ett minst lika extremt utfall som det observerade}) \\ &= P(T(X) \leq -2.45 \mid H_0 \text{ är sann}) = \Phi(-2.45) = 1 - \Phi(2.45) \approx 0.0071 < 0.01, \end{aligned}$$

så H_0 förkastas till förmån för H_1.

Svar: (a) $I_\sigma \approx (0.94, 3.71)$, (b) ja, det kan man.